Sheet ○ Slides

DONE BY

Rama Abbady

CONTRIBUTED IN THE SCIENTIFIC CORRECTION

Rama Abbady

CONTRIBUTED IN THE GRAMMATICAL CORRECTION

Rama Abbady

DOCTOR

Ma'moun Ahram

➢ Analysis of gene expression: RNA level

- New methods were discovered to measure the level of RNA molecules in a cell. They're used to analyze the expression of hundreds and thousands of genes.
- Remember! One way of detecting RNA molecules is **Northern Blotting** but it's used when the sample size is small (5-15).
- These methods start with the production of complementary DNA (**cDNA**).

❖ First! Make a complementary DNA (cDNA).

- cDNA: is a DNA molecule that's synthesized from RNA molecules.
- A viral enzyme called **Reverse Transcriptase** synthesizes cDNA from mRNA by a mechanism known as "Reverse Transcription".
- mRNA is used as a template to produce cDNA.
- The enzyme can't initiate the synthesis of cDNA by itself. It needs a primer.

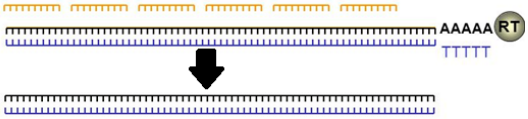
> **Remember!**
>
> -Transcription: DNA → mRNA.
>
> -Reverse Transcription: mRNA → cDNA.

- A primer needs to hybridize to complementary sequence on the mRNA molecule. Remember that during mRNA processing, a poly-A tail (AAAAA) is added to the mRNA which means that all mRNA molecules have a poly-A tail.
- Therefore, an oligo **dT** (deoxy-thymine) primer is used to hybridize to the poly-A tail which is found on all mRNA molecules.
- The synthesis of cDNA goes in the following steps:
  - The oligo dT primer hybridizes with the poly-A tail on the template mRNA.

  - Once the primer binds, Reverse Transcriptase extends the first cDNA strand.

  - The enzyme Reverse Transcriptase has an exonuclease activity. After generating the first cDNA strand, it degrades the template mRNA.

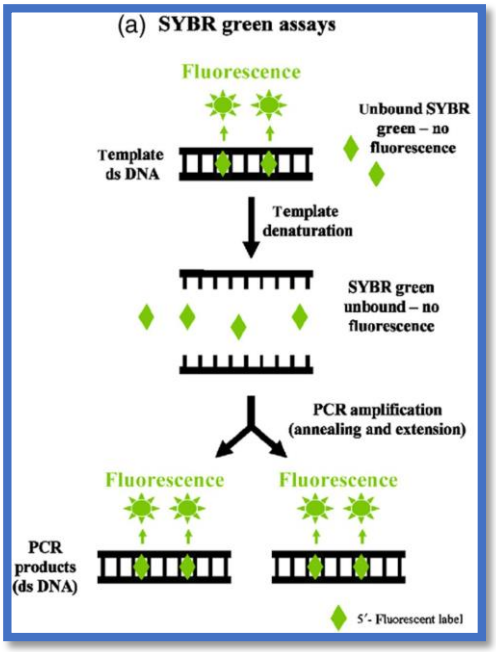
  - After digesting the template mRNA, the enzyme displaces it and synthesizes the seconds cDNA strand generating a double stranded cDNA molecule.



- **In summary**, Reverse Transcriptase has 3 functions:
  1. It synthesizes the 1st cDNA strand once the primer binds.
  2. It degrades the template mRNA.
  3. It synthesizes the 2nd cDNA strand → generating a double stranded cDNA.
- Since the template mRNA always gets digested, each mRNA molecule can produce only one cDNA molecule. (10mRNA → 10cDNA).

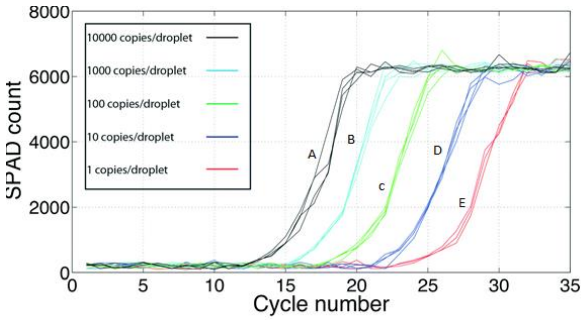
❖ Quantitative real-time PCR of mRNA

- One way of analyzing the level of gene expression and the quantity of mRNA molecules is quantitative real-time PCR.
- Since each mRNA gives only 1 cDNA, measuring the level of cDNA gives us an insight on the level of the mRNA and the expression of that gene (cDNA=mRNA).
- Quantitative real-time PCR: a method that's used to analyze **the quantity** of **mRNA** molecules by polymerase chain reaction in the presence of SYBR green.
- Remember! In PCR, DNA polymerase is used for repeated replication of a defined DNA molecule. The number of DNA molecules increases exponentially, doubling with each round of replication, amplifying small amounts of DNA so that it can be detected.



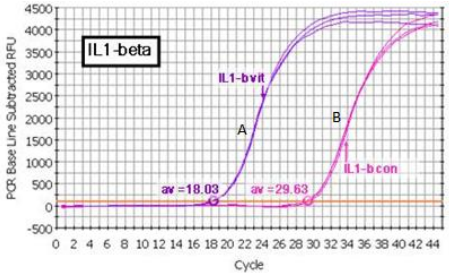
- SYBR green is a chemical (fluorescent) that can bind to double stranded DNA. When it binds, it emits light indicating the presence of the gene of interest.
- In quantitative real-time PCR, the starting material is cDNA.
- SYBR green binds to double stranded cDNA molecules synthesized from mRNA. The higher the amount of cDNA (mRNA), the earlier the light is emitted and the sooner it's detected.

- The following chart demonstrates the SYBR Green fluorescence produced in real-time PCR analyzing the expression of different genes in a cell.

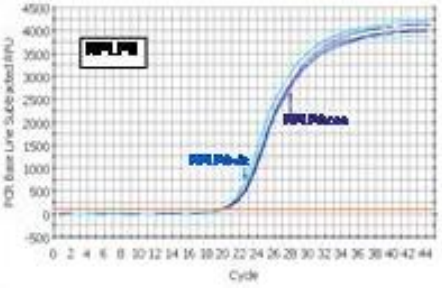


- According to the fact that the higher the amount of cDNA, the sooner the light is detected we can tell from the following chart that:
  - ✓ mRNA levels of gene **A** are the highest (it appeared **first**).
  - ✓ mRNA levels of gene **E** are the lowest (it appeared **last**).
  - ✓ Gene **B** is expressed at a higher level than gene **C**.
  - ✓ Gene **D** is expressed at a lower level than gene **C**.

- Quantitative real-time PCR can also be used to compare the expression of a certain gene in different cells under certain conditions.
- The expression of certain genes differs from one cell to another.
- This chart demonstrates the SYBR Green fluorescence produced in real-time PCR analyzing the expression of gene X in cell A and cell B.
- We can tell from this chart that gene X is expressed **more** in cell A than cell B.



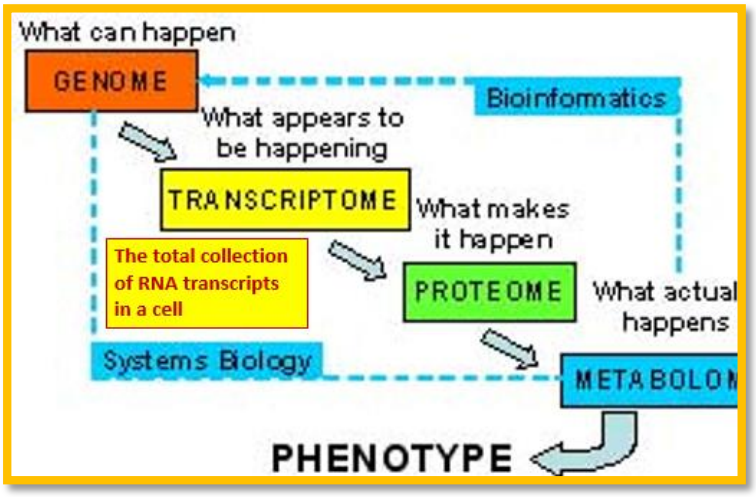
- In order to test the reliability of quantitative real-time PCR, scientists analyzed the expression of housekeeping genes.
- Housekeeping genes: genes that are expressed at all times at the same level in all cells. The expression of housekeeping genes isn't affected by the conditions of the cell. They have **unaltered expression.**
- This chart demonstrates the SYBR Green fluorescence produced in real-time PCR analyzing the expression of a housekeeping gene in two different cells.



- Notice how the two curves overlapped which means that the light was detected at the same time indicating that the gene is expressed at the same level in both cells.
- This shows the reliability of quantitative real-time PCR.

❖ The science of -omics

- -omics: the study of the total collection of something. It refers to a field of study in biology such as gen**omics**, transcript**omics**, prote**omics** or metabol**omics**.
- Gen**omics**: the study of the whole genome in a cell.
- Transcript**omics**: the study of the total collection of RNA transcripts in a cell.
- Prote**omics**: the study of the total collection of proteins in a cell.
- Metabol**omics**: the study of the total collection of metabolites (vitamins, carbohydrates...) present in a cell.
- Phospho-prote**omics**: the study of the total collection of phosphoproteins in a cell.



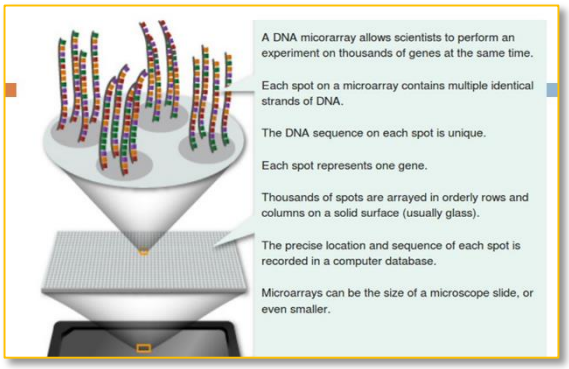
❖ Studying the transcriptome

- Studying all the expressed genes in a cell simultaneously.
- Remember that not all genes are expressed. Gene expression is variable from one cell to another.
- By examining the expression of so many genes simultaneously, we can understand gene expression patterns in physiological and pathological states.
- Remember that RNA molecules undergo splicing, which means that one gene can give more than one RNA molecule. So the number of RNA molecule is larger than the number of expressed genes.
- One such method in studying transcriptomes is DNA microarrays.

❖ DNA microarrays

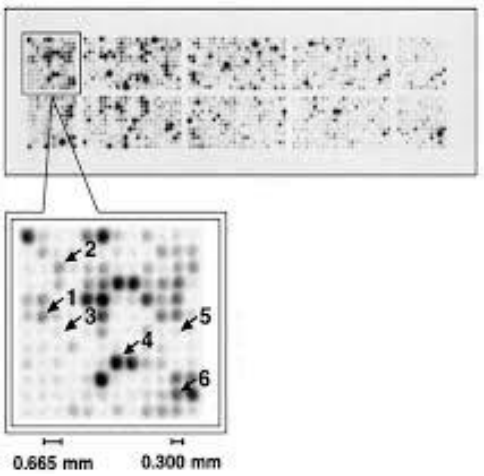
- It is a technology that allows us to study the expressed genes in a cell at certain conditions simultaneously. It's used to analyze the expression of hundreds and thousands of **known** genes.
- DNA microarrays are glass microscope slides spotted with up to tens of thousands of DNA fragments in an area the size of a fingernail.
- It consists of a glass chip onto which probes are printed in small spots at high density.
- Each spot on the array consists of a single type of probe → each spot represents one gene.
- Tens of thousands of unique probes can be printed onto the chip, so it's possible to produce DNA microarrays containing sequences representing all of the genes in cellular genome.
- The exact sequence and position of every DNA fragment on the array is known. For example, a probe for gene X is always found at position n.1 on the array, if a strong signal is emitted from that position, we know that gene X is expressed in this cell at high levels.

> Remember! A probe is an oligonucleotide (short DNA) that's used to detect DNA molecules with specific sequences by hybridizing to them.



- DNA microarrays starts with the production of complementary DNA (**cDNA**).
- First, we extract mRNA molecules from the cell.
- mRNA molecules are then converted to cDNA molecules by Reverse Transcriptase. Since each mRNA gives only 1 cDNA, the amount of cDNA molecules reflects the amount of mRNA in a cell.
- cDNA molecules are then labeled by radioactive probe.
- We add the radioactive cDNA to the array. cDNA molecules will hybridize to their complementary probes on a certain position on the array.
- If a gene is expressed, then the cDNA will exist and bind to a specific complementary DNA fragment (probe) on the microarray.
- Since the cDNA is labeled, binding can be detected and the expression of this gene is determined.

- Each spot (position) represents one specific gene. When a signal is emitted from a certain spot, it indicates that this gene in expressed.
- Let's say that the probes at positions 2, 3, 4 can hybridize to the cDNA of the genes X, Y, Z respectively.
- At position n.4, there's a strong signal which means there's high binding of the probe in that position with the cDNA → gene **Z** is expressed at high levels in this cell.



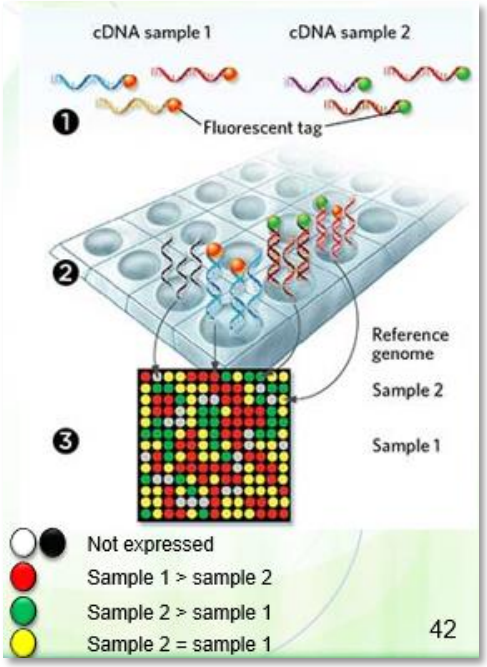
- At position n.2, the signal is faint. There's little amount of cDNA → gene **X** is expressed in this cell but at low levels.
- At position n.3, there's no signal. There's no cDNA → gene **Y** is silent in this cell (not expressed).
- This is done for a single sample using radioactively labeled cDNA.

❖ DNA microarrays: comparative gene expression

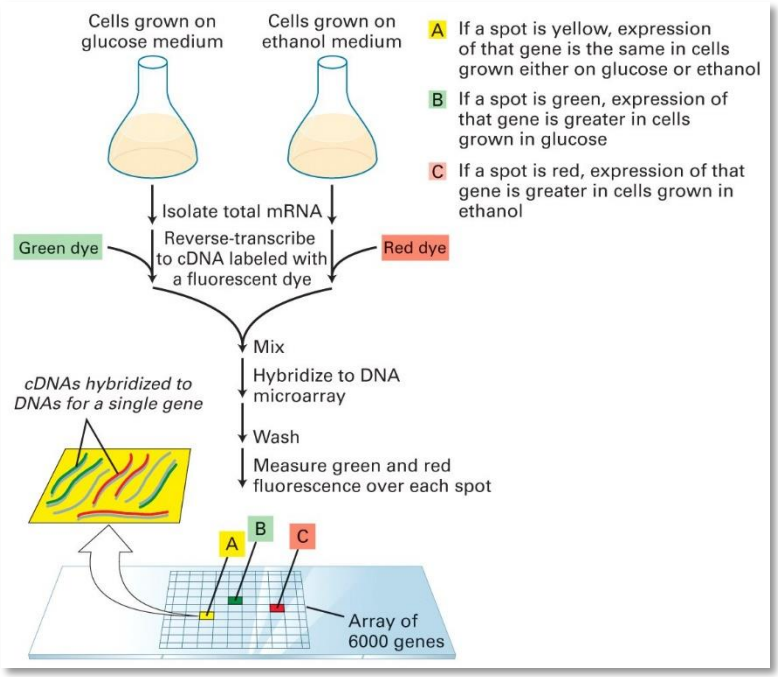
- DNA microarrays can also be used to compare the expression of all genes in different samples.
- We can use it to compare the expression of genes in a certain cell when it's exposed to different conditions (heat, ethanol, glucose, a certain drug...), for example: comparing gene expression between a normal cell and a cancer cell.
- First we extract the mRNA molecules from both samples (normal, non-treated cell and the treated cell) and we convert them to cDNA molecules.
- In comparative gene expression, cDNA molecules are labeled by a fluorescent not radioactively.
- Let's say for example that we want to compare the expression of genes in a cell before and after treating it with a drug. We add for sample n.1, the normal cell, a red fluorescent and we add for sample n.2, the treated cell, a green fluorescent.
- We then combine the two samples and add them to the array. The red and green cDNAs will start competing over binding with the probes on the array.
- An increase in the amount of a RNA molecule in one sample versus the other is reflected by an increase the amount of produced cDNA and an increase in fluorescence in the bound spot.

- Now there are 4 possibilities:

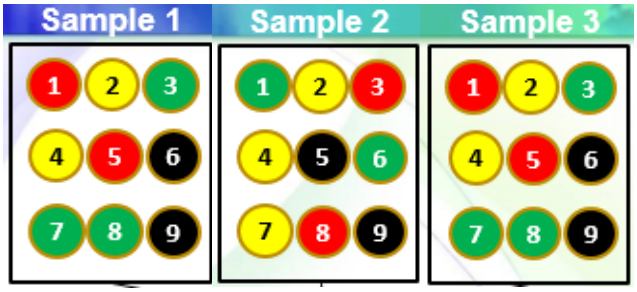
1. No effect! If the expression of genes in the normal cell is equal to the expression of genes in the treated cell, a **yellow** signal will be emitted. The number of red **cDNA** will be equal to the number of green **cDNA**, which means they'll have equal chance of binding. (**red** + **green** = **yellow**).
2. No effect! If the gene is silent under both conditions, there'll be no red **cDNA**, no **green** cDNA, no signal will be emitted (**Black** or White).
3. If the gene expression becomes more active after treating the cell, there'll be more green **cDNA** than red **cDNA** which means that the green **cDNA** will have higher chance of binding to the probes → a **green** signal will be emitted.
4. If the gene expression becomes less active after treating the cell, there'll be more red **cDNA** than green **cDNA** which means that the red **cDNA** will have higher chance of binding to the probes → a **red** signal will be emitted.



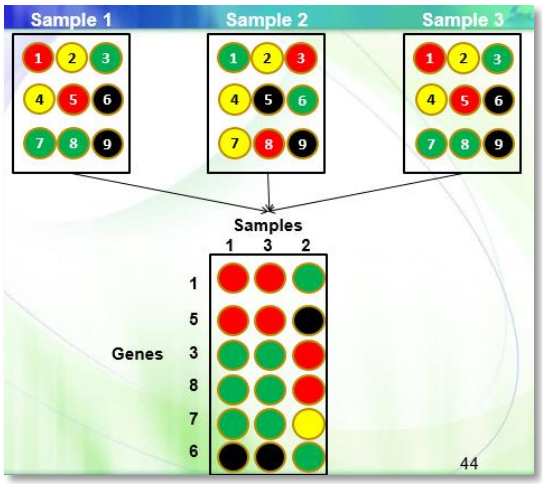
- Let's do another experiment. We want to grow sample n.1 on ethanol medium and label the **cDNA** with a red fluorescent, and grow sample n.2 on glucose medium and label the **cDNA** with a green fluorescent to compare gene expression under both conditions.



- This can be done at a larger scale whereby samples are compared to the same control sample and a computer program combines all data illustrating differences in expression among the samples and classifying them into different groups.
- In this experiment, we exposed three samples to three different conditions and then we compared the expression of 9 genes of each sample with the same control sample.
- A Control Sample: is a normal/healthy/untreated sample that we compare other samples to.
- All three samples are labeled with red fluorescent and the control sample is labeled with green fluorescent.

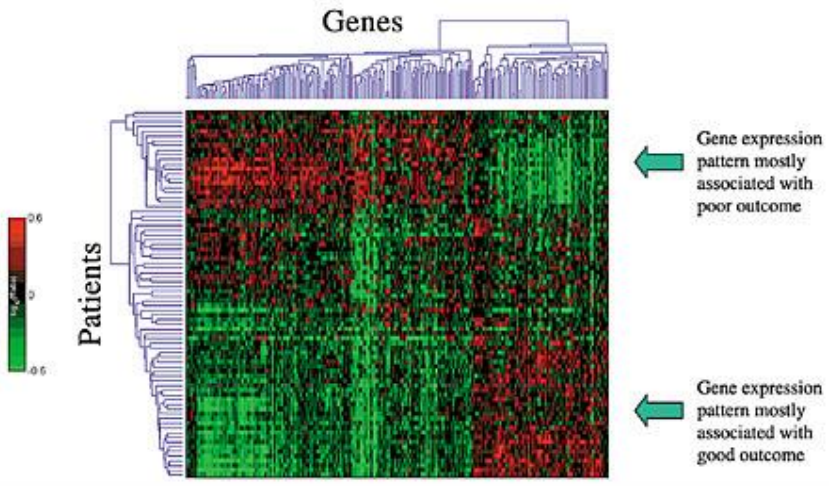


- Now let's analyze the results:
  - ✓ Gene n.1 has higher expression in samples 1,3 than in the control sample → **RED.**
  - ✓ Gene n.1 has higher expression in the control sample than in sample 2 → **green.**
  - ✓ There's equal expression of gene n.2 in samples 1,2,3 and the control sample → **Yellow** in all results.
  - ✓ Gene n.9 is silent in all samples (1,2,3 and the control sample) → **Black**.
  - ✓ There's higher expression of gene n.3 in the control sample than in sample 1,3 → **green**.
  - ✓ There's higher expression of gene n.3 in sample 2 than in the control sample → **red**.
  - After analyzing the gene expression in all samples, the computer eliminates the results of the genes that are not informative (same result in all samples). In this experiment: genes 2, 4 and 9.
  - The computer then combines the results and clusters samples according to expression ***(Sample 1 and 3 have similar results, while sample 2 is different).***



❖ DNA microarrays and clinical application

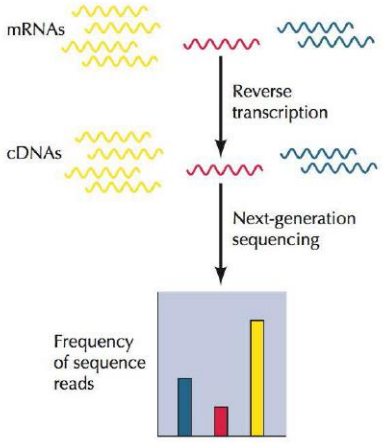
- Scientists have managed to successfully associate clinical diagnosis with gene expression.
- According to gene expression, by taking a biopsy of a tumor in a cancer patient we can tell whether this tumor is malignant or benign and whether this patient will respond to therapy or not and how well that response will be.
- First, doctors took tumor samples from both cancer patients who respond to therapy (sample 1) and who don't respond to therapy (sample 2). They extracted mRNA molecules from both samples to analyze gene expression using DNA microarrays.
- They labeled cDNA in sample 1 and 2 with red fluorescent and compared them to a control sample (healthy sample) that's labeled with green fluorescent. Then they analyzed the results:
  - ✓ If the signal is green → this gene has higher expression in normal healthy cells than in cancer cells (under expressed in cancer cells-less active).
  - ✓ If the signal is red → this gene has higher expression in cancer cells than in normal healthy cells (overexpressed in cancer cells-more active).
- Note that genes that are silent or have equal expression in all samples are considered to be not informative and are always eliminated from the results.
- The computer clusters the samples, combining similar results beside each other.
- After analyzing the results, scientists noticed that patients with a certain gene expression pattern respond to therapy and patients who have a different gene expression pattern don't.
- Analyzing gene expression can also help doctors decide the effective dose for each patient. Doctors can also tell whether this patient will have a good outcome or a poor outcome to therapy and according to that the doctor gives extensive therapy to patients with poor outcome and light therapy to patients with good outcome.



❖ RNA sequencing (RNA-seq)

- DNA microarrays is limited to detect transcripts corresponding to **known** genomic sequences. It is used to detect mRNA molecules that we **know** the sequence of. We have to design specific probes that can hybridize to the cDNA of interest.
- RNA sequencing is used to determine and quantify all of the RNA molecules expressed in a cell including (mRNA, tRNA, micro-RNA,etc.).
- In contrast to microarray analysis, RNA-seq reveals the complete extent of transcribed sequences (RNAs) in a cell, rather than just detecting those that hybridize to a probe on a microarray.
- This analysis determines the **abundance** (quantity) as well as the **identity** of all transcribed sequences (known and unknown).

- RNA sequencing goes in the following steps:
  - We extract RNA molecules from the cell.
  - We convert RNA molecules to cDNA by reverse transcription.
  - We do next generation sequencing to reveal the entire sequence of cDNA molecules.



- The relative amount of each cDNA (mRNA) is indicated by the **frequency** at which its sequence is represented in the **total** number of sequences read.

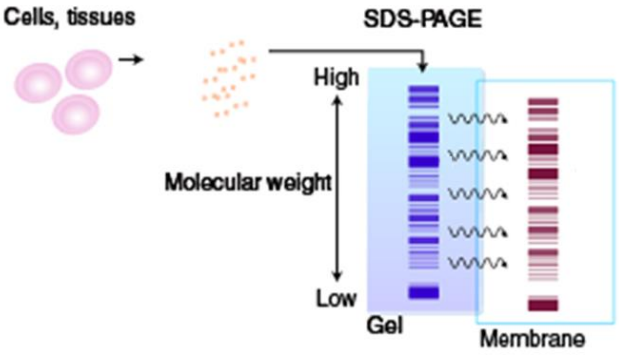
- RNA-seq can be used to:
  - ✓ Characterize and detect novel transcripts (unknown RNAs).
  - ✓ Identify splicing variants.
  - ✓ Profile the expression levels of known transcripts.

- RNA-seq has revealed that many more RNAs are transcribed than are translated. It has identified coding RNAs as well as non-coding RNAs.

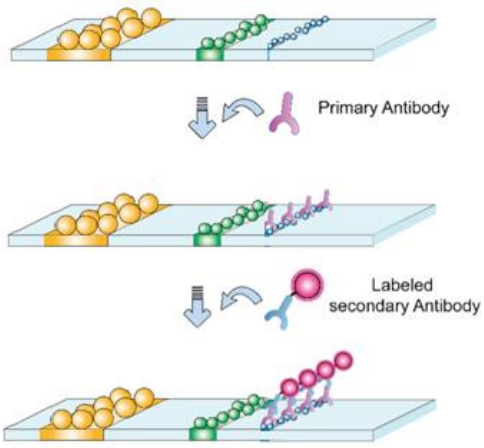
➢ Analysis of gene expression: protein level

❖ Immunoblotting (Western Blotting)

- Western blotting is a technique used for the analysis of individual proteins in a protein mixture. It's used to identify and quantify proteins in cells.
- Immunoblotting goes in the following steps:
  - Proteins are collected from cells then separated by SDS-PAGE.
  - After separating them, we transfer proteins from the gel to a membrane (similar to northern and southern blotting).



  - We add a primary antibody that can bind to the target protein specifically.
  - Then we add a secondary antibody that targets the primary antibody.
  - Once the secondary antibody binds to the primary antibody it generates a signal indicating the presence of the protein. Depending on the intensity of the signal we can determine the quantity of the protein in the cell.



- What information do we get?
  - ✓ Detect the presence of the protein. **(If there's a signal).**
  - ✓ Determine the quantity of the protein. **(depending on the intensity of the signal).**
  - ✓ Determine the molecular weight (size) of the protein**. (SDS-PAGE).**